

# Weekly Report

2012.10.1-2012.10.8

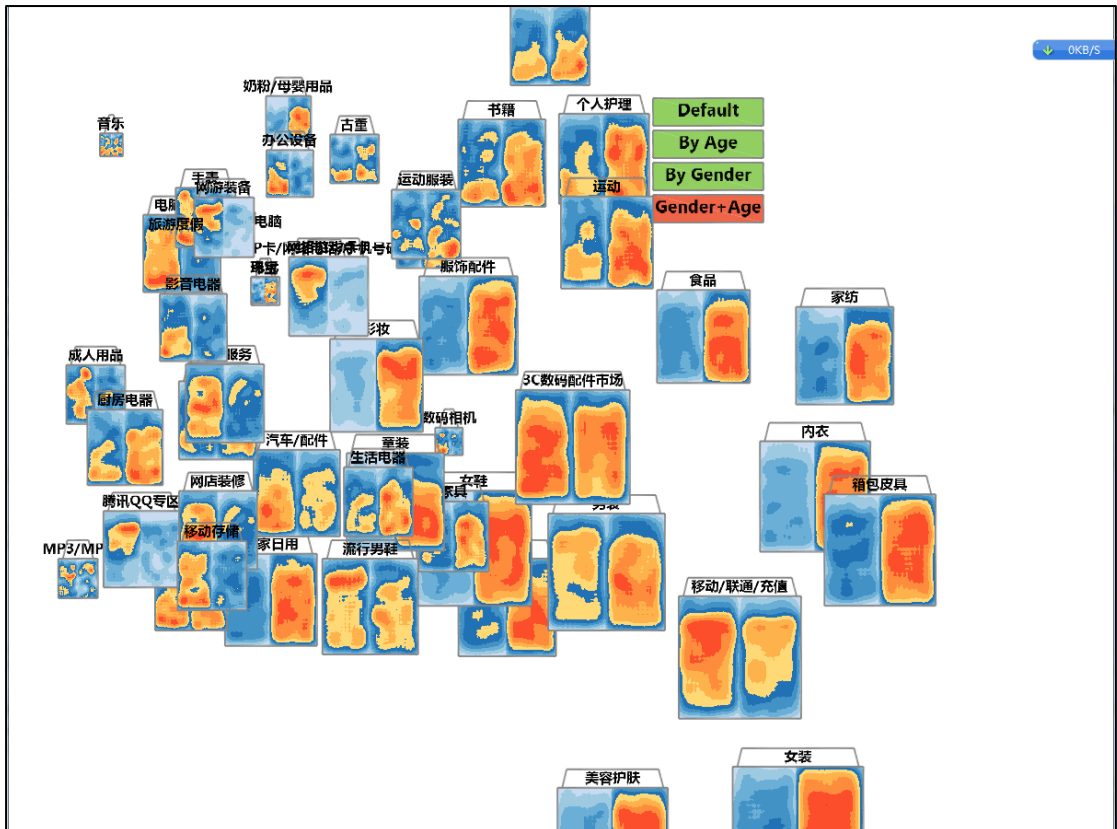
黄芯芯

## 本周工作：

（上周因为上课、VINCI 会议等，没有太多工作进展，因此把两周的内容合在一起。）

### 1. 淘宝标签项目：

- 1) 根据淘宝提供的类目之间的关系数据，将类目用 MDS 投影，并且对之前重叠的问题作了一些处理，每个矩形的面积按照购买过该类目人数进行编码，得到如下的结果，效果不好。



- 2) 后来去“淘宝指数”上看，淘宝指数所做的也是展示购买或搜索某个类目的人群特征，但是他们只是简单地在单个标签（按照性别、年龄、地域等）展示人群特征，类似于信息图，而我们需要做的正是综合多个标签，因此需要找到一个可扩展到多个标签的展示方式。所以我尝试做了下面一个 sunburst 图，主要是从 sunburst 的可扩展性出发，觉得可以用每一层代表一个标签，每添加一个标签属性，就相当于在每个类目上多加一层。

下图中，里层是性别属性，浅蓝色代表女性用户，浅绿色代表男性用户；外层是年龄属性，不同颜色代表不同年龄段，外层展示了分别在男女性别上的年龄分布。如果再添加属性，则继续在外层进行添加。初始距离用 MDS 进行投影，对重叠的问题稍微用力引导做了调整。相对于原来的结果，这个 sunburst 好像看不出来什么 pattern，但是能不能通过交互或者改变每层的排布来改善。

看屈老师他们做的 DICON，其实个人感觉他们只是对聚类的 merge 和 split 的操作做的比较好，以及可以嵌入到其他的可视化形式中，比如地图、散点图等。所以我也想，



下周计划:

1. 周三组会报告。
2. 明确标签项目的下一步计划。
3. 上课&课程作业。
4. 写教材。